Module 5 Lab Activity: Working with Categorical Variables
PSY 652 Research Methods
Sep 30, 2020

This activity is designed to build familiarity with categorical variables and categorical analyses in R. In this activity you will create a new R notebook using the slp_cat.csv dataset.

The following files may be found in the "Module 5 Lab" section of your modules section:
   Dataframe: slp_cat.csv

1. Create a new R notebook and name it "Category_variables_notebook"
2. Create a new R chunk with a first level header: "Load Libraries"
    a. load the tidyverse & epitools package in this R chunk (You may need to install the epitools package)
3. Create a first level header called: "Import Data"
    a. Insert a new r-chunk and read in the "sleep_cat.csv" dataset and assign it to an object named "slp"
4. Create first level header called: "See the structure of the data frame"
    a. Insert an R chunk and use the str() function to view the structure of the slp dataframe
5. Create a first level header called: "Get a summary of the dataset"
    a. Insert an R chunk and use the summary() function to summarize your dataset
    b. In the white space below write, "My variables are in character format and I need to convert them to factors"
6. Create a first level header called: "Factor variables via mutate()"
    a. Use the mutate() function to factor the *sex, cond, prior*, and *sleep* variables. Save them back into your slp dataset.
        i   Hint: dataset <- mutate(dataset, variable = as.factor(variable))
7. Create a first level header called "See the structure of the dataset again"
    a. Insert an R chunk and use the str() function to see the new structure of your dataframe object
8. Create a first level header called: "Get a summary of the dataset again"
    a. Insert an R chunk and use the summary() function to summarize your dataset
9. Create a first level header called: "Run a chi-square of sex and sleep variables"
10. Create a 2nd level header called: "Step 1: Create a frequency table of your variable"
    a. Insert an R chunk and create a table object named "sex_sleep_table" with sex levels as the rows, and sleep levels as the columns.
        i   Hint: First select() the variables of interest in a new dataframe object, then use the table() function to table them into a new object called sex_sleep_table.

Your new frequency table object should look like this:

```
sex_sleep_table
```

```
          sleep
sex        Bad_sleep  Good_sleep
  Female          67         180
  Male           230         123
```

11. Create a second level header called: "Step 2: run chisq.test() on the new frequency table object"
    a. Insert a new R chunk and run a chi-square test on your sex_sleep_table object
    b. In the white space below, interpret the output
12. Create a 3rd level header: "Step 2.1: Calculate the contingency coefficient effect size"
    a. Insert a new R chunk and calculate the contingency coefficient effect size
        i Hint: equation = sqrt(X-squared / (X-squared + N))
    b. In the white space below, tell me what this effect size means in one sentence.
13. Create a first level header called: "Get odds ratio of our sex_sleep_table"
14. Create a second level header called: "Calculate it by hand"
    a. Using the values from your sex_sleep_table, calculate the odds ratios of males having good_sleep as compared to females.
        i   You may use any method to calculate the Odds ratio (As in, you don't need to use R), but please provide me the answer somewhere in the white space.
15. Create a second level header called: "Calculate it via oddsratio.wald() from the epitools package"
    a. Insert a new R chunk and use the oddsratio.wald() function to get the odds ratio of males having good_sleep as compared to females
        i   Hint: oddsratio.wald(table_object)
        ii   This value should match the value you calculated in 14a.

16. Once you've completed all of these steps, Restart R and Run All Chunks, and then preview your notebook. Save your notebook as both a .Rmd and an html file and exit RStudio.

17. Upload the html version of your notebook to the assignment called "Turn in Module 5 Lab Activity here" on the course Canvas page.